BOSTON COLLEGE
Department of Economics
EC771: Econometrics
Spring 2009
Prof. Baum, Ms. Skira

PROBLEM SET 4: DUE TUESDAY 24 MARCH 2009 AT CLASSTIME

1. Using the "canned" dataset `earnings`, which you may access from within Stata with the command

```
.use http://fmwww.bc.edu/ec-p/data/dmackinnon/earnings
```

Run the regression

$$y_t = \beta_1 d_{1t} + \beta_2 d_{2t} + \beta_3 d_{3t} + u_t$$

Where the d_i are dummy variables and y is `earnings`. Then test the null hypothesis that $E(u_t^2) = \sigma^2$ against the alternative that

$$E(u_t^2) = \gamma_1 d_{1t} + \gamma_2 d_{2t} + \gamma_3 d_{3t}$$

Using both an F test and Stata's `robvar` command. Interpret.

2. If u_t follows the stationary $AR(1)$ process

$$u_t = \rho u_{t-1} + \epsilon_t,\ \epsilon_t \sim IID(0, \sigma_\epsilon^2),\ |\rho| < 1$$

show that $Cov(u_t u_{t-j}) = Cov(u_t u_{t+j}) = \rho^j \sigma_\epsilon^2/(1-\rho^2)$. Then use this result to show that $Corr(u_t u_{t-j}) = \rho^j$.

3. Consider the model

$$c_t = \alpha + \beta c_{t-1} + \gamma_0 y_t + \gamma_1 y_{t-1} + \epsilon_t$$

where c and y are the logarithms of consumption and income, respectively. Show that this model contains as a special case the following linear model with $AR(1)$ errors;

$$c_t = \delta_0 + \delta_1 y_t + u_t,\ u_t = \rho u_{t-1} + \epsilon_t$$

where ϵ_t is IID. Write down the relation between the parameters δ_0, δ_1, ρ of this model and the parameters $\alpha, \beta, \gamma_0, \gamma_1$ of the former model. How many restrictions are imposed on the latter set of parameters by the second model?

4. Using the "canned" dataset `consumption`, which you may access from within Stata with the command

```
.use http://fmwww.bc.edu/ec-p/data/dmackinnon/consumption
```

a) estimate the nonlinear model that is implicitly defined by the model of (3) with $AR(1)$ errors for the period 1953q1–1996q4 by nonlinear least squares (`nl` in Stata). Do not use a specialized procedure for $AR(1)$ estimation. For starting values, use the estimates of δ_0, δ_1 implied by the OLS estimates of

$$c_t = \alpha + \beta c_{t-1} + \gamma_0 y_t + \gamma_1 y_{t-1} + \epsilon_t$$

What are the properties of these OLS estimates?
b) Test the restrictions that the nonlinear model imposes on this linear model by means of an asymptotic F test.

5. Using the "canned" dataset `discrim`, which you may access from within Stata with the command

```
.use http://fmwww.bc.edu/ec-p/data/wooldridge/discrim
```

(a) How many of the observations are from Pennsylvania? (hint: `describe` may be useful)
(b) What is the average starting wage (first wave)? What are the min and max of this variable?
(c) Test the hypothesis that average incomes in NJ and PA are equal. (hint: `help ttest`)
(d) Test the hypothesis that the average price of an entree (first wave) was $1.39.
(e) Write a Stata program, using the `ml` syntax, which will estimate the parameters of a k-variable linear regression model with a constant term, including the σ^2 parameter, via maximum likelihood. Use your program to estimate the model:

$$pfries_i = \beta_0 + \beta_1 income_i + \beta_2 prpblck_i + \epsilon_i$$

over the whole sample, and
(f) only over the New Jersey observations (those for which `state` equals 1).
(g) Use Stata's `regress` command to estimate the same linear regression, and use `bootstrap` to generate bootstrap standard errors for the $\hat{\beta}$ parameters. Discuss how the bootstrap confidence intervals for `income` and `prpblk` compare with the conventional confidence intervals computed by `regress`.